



Webb, MW., Hunukumbure, MR., Beach, MA., & Nix, AR. (2006). The effect of computation and feedback delay on the capacity of multiuser MIMO systems in a small outdoor cell. *IEEE 63rd Vehicular Technology Conference, 2006 (VTC 2006-Spring)*, 4, 1988 - 1992. <https://doi.org/10.1109/VETECS.2006.1683194>

Peer reviewed version

Link to published version (if available):  
[10.1109/VETECS.2006.1683194](https://doi.org/10.1109/VETECS.2006.1683194)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# The Effect of Computation and Feedback Delay on the Capacity of Multiuser MIMO Systems in a Small Outdoor Cell

Matthew Webb, Mythri Hunukumbure, Mark Beach and Andrew Nix

Department of Electrical and Electronic Engineering

University of Bristol, Bristol, BS8 1UB, UK

Email: {M.W.Webb, Mythri.Hunukumbure, M.A.Beach, Andy.Nix}@bristol.ac.uk

**Abstract**—This paper examines how the capacity of three multiuser multiple-input multiple-output (MIMO) algorithms is affected by a delay in computing and feeding back the transmit (and receive) weighting matrices used by the algorithms. Iterative schemes determining transmit weights that achieve the system-wide Nash Equilibrium (NE) or block-diagonalization (BD) of the overall system channel matrix are compared as is a scheme applying successive diagonalization (SD) among users. The NE is found to be far the most robust to out-of-date channel information, whilst SD is found to suffer badly at even the shortest delay considered here (about 0.9ms); BD's losses vary over a wide range depending on the delay. For the two iterative algorithms we also examine the tradeoff between better convergence and less delay. In both cases, the loss in capacity for weaker convergence is very small, and so it is desirable to minimize the delay and thus the capacity loss caused by it.

## I. INTRODUCTION

The information-theoretic capacity benefits of employing dual antenna arrays in future wireless systems, to construct a multiple-input multiple-output (MIMO) system, are now widely appreciated [1]. Considerable work has been undertaken to further the understanding of MIMO systems in a single-user scenario, but means of exploiting the multi-user MIMO channel have been less widely investigated.

In earlier work [2], we have examined the behaviours of three multi-user access algorithms: one based on the non-cooperative Nash equilibrium [3] and two based on transmit-side orthogonality [4]–[7]. This work made the common, but unrealistic assumption that there was no delay between measuring the channel and applying the transmit/receive weights. In reality, there will be both a computation and a feedback delay, particularly since two of the algorithms are iterative, so that the channel will have changed by the time the weights calculated on the basis of it are applied. This paper investigates the effect of such application delays on capacity, characterizes the losses incurred and examines the tradeoffs that may be made. Our studies are based on a multi-user small, urban cell uplink scenario and take advantage of measured data obtained in a UK city centre. We describe the system model we will use and the manner in which the measured data were obtained and utilized in Section II. Section III gives a brief description of the three algorithms and Section IV presents results examining how delay affects the capacity of these systems and what

tradeoffs can be made between better convergence and shorter delay. Section V gives our conclusions.

## II. SYSTEM MODEL AND MEASUREMENTS

We consider a quasi-static, narrowband uplink with multiple users communicating simultaneously with a single base station (BS) equipped with  $n_R$  antennas. We will equip all  $L$  users identically, with  $n_T$  transmit antennas and the same mean signal-to-noise ratio (SNR) at the BS. The  $j^{\text{th}}$  user will use a transmit weighting matrix  $\mathbf{T}_j$  and their signal will be extracted by the BS with  $\mathbf{R}_j$ . We thus model  $\mathbf{y}_j$ , the weighted received signal for user  $j$  at a point in time, as:

$$\mathbf{y}_j = \mathbf{R}_j \left( \sum_{i=1}^L \mathbf{H}_i \mathbf{T}_i \mathbf{x}_i + \mathbf{w} \right). \quad (1)$$

Here,  $\mathbf{H}_i$  is the channel matrix from user  $i$  to the BS,  $\mathbf{w}$  is unit-variance, zero-mean, complex additive white Gaussian noise (AWGN) and  $\mathbf{x}_i$  is a data vector of arbitrary dimension.

### A. Measurement Setup

The measured data were obtained [8] in the city-centre of Bristol, UK, an area with dense urban clutter. At the transmit side, two dual-polarized, 65° bandwidth UMTS antennas with 20-wavelength separation were roof-mounted on a five-storey building about 30m high which is on top of a hill overlooking the city centre, and were down-tilted by 8° from the horizontal. For the receiver, a uniform circular array (UCA) of 8 monopoles with half-wavelength radial spacing was used. This was vehicle mounted and driven around the area shown in Fig. 1. At each of 24 positions, 1024 MIMO snapshots were recorded, each well within the channel coherence time, in 8 blocks of 1.3s duration with 1.5s recording time between blocks. The measurements were taken at 128 distinct frequency fingers spanning 20MHz centered on 1.92GHz. In this work, we will average each snapshot's capacity over all 128 narrowband frequencies.

Since this setup has a *single* transmitter and *many* receivers we will transpose the channel matrices and view each measurement as a transmission from one of the 24 locations back to the receiver.

When in use in the simulations of Section IV, the channel coefficients were normalized using the data from the full

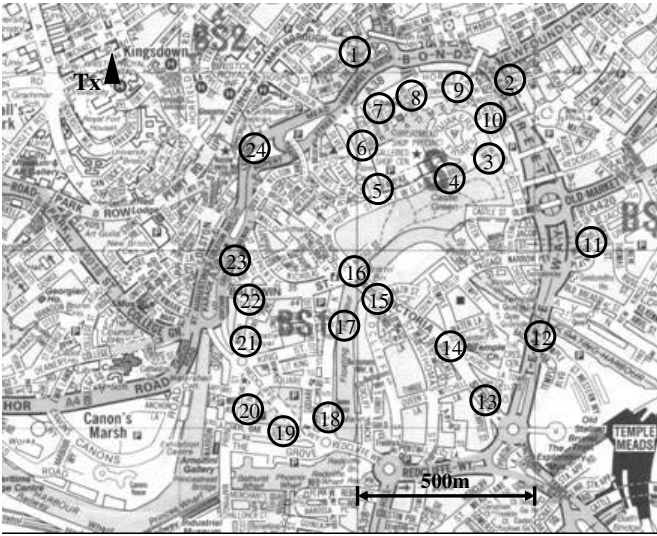


Fig. 1. Measurement drive locations and transmit site, superimposed on a map of Bristol city-centre, UK, on a 500m<sup>2</sup> grid.

20MHz bandwidth to remove the mean path-loss at each snapshot. Denoting the coefficient at frequency  $f$  between transmit antenna  $T$  and receive antenna  $R$  as  $g_{f,T,R}$ , the normalized coefficients are:

$$h_{f,T,R} = \frac{g_{f,T,R}}{\sqrt{\left(\frac{1}{128n_T n_R}\right) \sum_{f=1}^{128} \sum_{R=1}^{n_R} \sum_{T=1}^{n_T} |g_{f,T,R}|^2}}. \quad (2)$$

We will only be using 4 antennas at each end of the link, and we will select transmit antennas no. 1–4.

### B. Interpolating the data

The aim this paper is to investigate the effect of the channel having evolved by the time the transmit/receive filters are actually applied. However, the measured data has a time-resolution of about 10ms/snapshot within each of the eight blocks. This is likely to be somewhat larger than a likely iterative and feedback delay of, say, a few milliseconds. We will thus linearly interpolate the measured data to increase its temporal resolution by adding ten interpolated snapshots between every pair of measured snapshots. This will increase the temporal resolution of the data to 0.9ms/snapshot. There will be one ‘bad’ snapshot at the end of each block before the 1.5s data-recording time, which we will discard. This interpolation is reasonable since, according to [8], each snapshot is recorded well within the channel coherence time and there is thus reasonable correlation between one snapshot and the next, with the exception of the discarded snapshots at the end of each measurement block.

## III. THE ALGORITHMS

We will examine three algorithms and here give brief descriptions of how they function. These algorithms all require either interference covariance knowledge or full channel knowledge of other users. We will therefore assume that the

base station is able to obtain this information, carry out the necessary calculations and feed back the results to the users.

### A. The non-cooperative Nash equilibrium (NE)

This [3] is a form of generalized waterfilling in which each user competitively tries to maximize their own capacity. We define the noise-plus-interference terms in (1) (i.e.  $w$  and all signal contributions for users  $i \neq j$ ) to have a covariance matrix  $\mathbf{Q}_j$  from user  $j$ ’s perspective. It has been shown [9] that to maximize their own capacity, each user should apply waterfilling on the ‘prewhitened’ channel  $\mathbf{Q}_j^{-1/2} \mathbf{H}_j$ . However,  $\mathbf{Q}_j$  will depend on what waterfilling matrix each other user has chosen. Thus, each user in turn selects their transmit matrix  $\mathbf{T}_j$  according to the waterfilling criterion derived from the current transmit schemes of the other users and we iterate among the users until the change in the total system capacity is less than some threshold. The Nash equilibrium results in that transmission profile from which no user would seek to deviate because to do so would, once all the other users have responded, result in a decrease in their individual capacity.

### B. Block diagonalization (BD)

This algorithm was originally designed in [4], [6], [7] (and in partial form in [5], [10]) to allow a base station to communicate with multiple users. We will here describe it in its original form, and then ‘transpose’ the algorithm to suit our scenario.

The BD algorithm seeks the zero-forcing solution to eliminate inter-user interference whilst allowing streams from one user to interfere. That is, it determines  $\mathbf{T}_j$  and  $\mathbf{R}_j$  such that

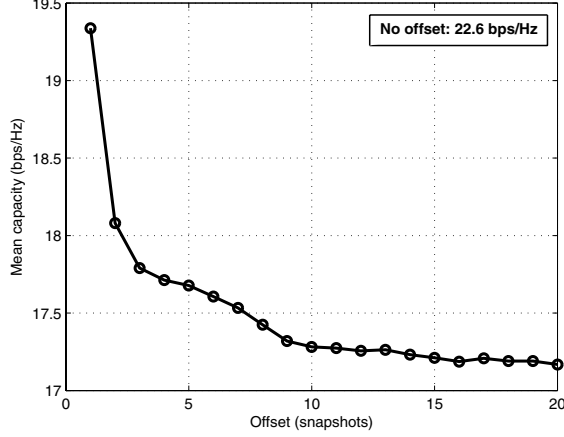
$$\mathbf{R}_j^\dagger \mathbf{H}_j [\mathbf{T}_1 \mathbf{T}_2 \cdots \mathbf{T}_L] = [\mathbf{0}_1 \cdots \mathbf{0}_{j-1} \mathbf{A}_j \mathbf{0}_{j+1} \cdots \mathbf{0}_L], \quad (3)$$

which is block-diagonal with  $\mathbf{H}_j$  temporarily the channel from the base to user  $j$ . This constitutes transmission from the single array at the base via a set of different  $\mathbf{T}_j$  and each user applying their own  $\mathbf{R}_j$  to maximize their end-to-end channel gains (in the main diagonal of  $\mathbf{A}_j$ ). This results in choosing the left- and right-nullspaces of the aggregated other-users’ end-to-end channels for  $\mathbf{T}_j$  and  $\mathbf{R}_j$  respectively, waterfilling over the eigenmodes at the transmitter and iterating until (3) is satisfied to within some threshold on the zero-elements.

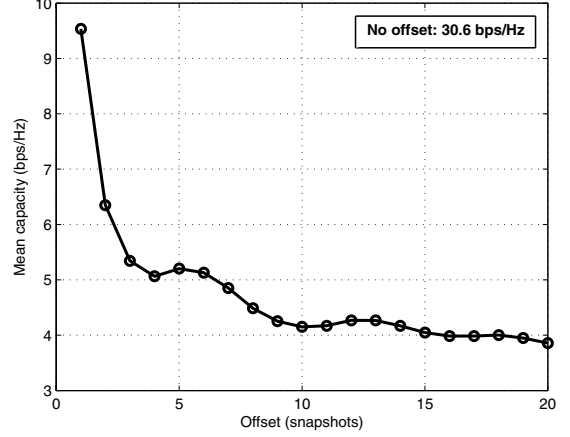
To adapt this system to our multi-point to point scenario, we will form a ‘virtual’ BS by aggregating the original users and view the original base station as multiple users separated by different channels. We must then transpose the form of (3) giving  $(\mathbf{R}_j^\dagger \mathbf{H}_j \mathbf{T}_j)^T = \mathbf{T}_j^T \mathbf{H}_j^T \mathbf{R}_j^*$ , which also recalls from Section II-A that we are transposing our channel matrices. Thus we transmit with  $\mathbf{R}_j^*$  and receive with  $\mathbf{T}_j^T$ .

As with e.g. BLAST [11], there are limits on the number of data streams that can be extracted from such a system. It can be shown [7] that, if  $N_j$  is the number of streams transmitted by user  $j$ , then we require

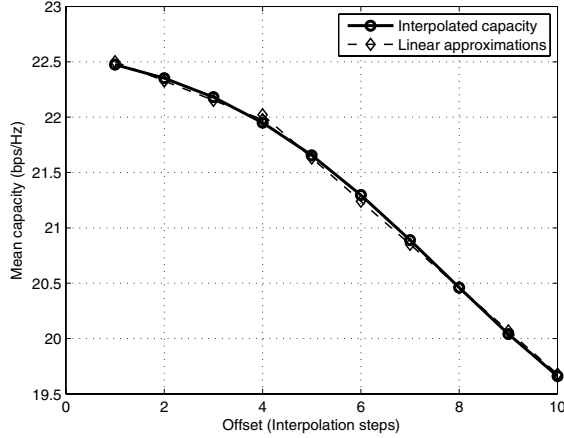
$$n_R \geq \sum_{j=1}^L N_j \text{ and } n_{T_j} \geq N_j \text{ and } L \leq n_R. \quad (4)$$



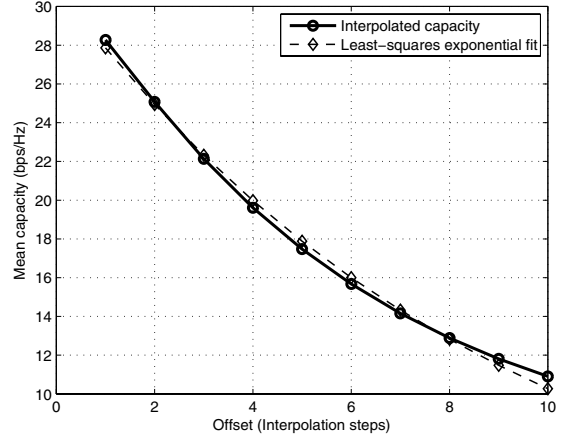
(a) Snapshot-by-snapshot capacity.



(a) Snapshot-by-snapshot capacity.



(b) Capacity at interpolated points between snapshots. Fitted curves from (5) also shown.



(b) Capacity at interpolated points between snapshots. Fitted curve from (6) also shown.

Fig. 2. Capacity of NE under channel offset. SNR = 20dB,  $n_T = n_R = 4$  for all users. Capacity with no offset is 22.6bps/Hz.

Fig. 3. Capacity of BD under channel offset. SNR = 20dB,  $n_T = n_R = 4$  for all users.

### C. Successive diagonalization (SD)

In this algorithm [5], [6], instead of requiring complete orthogonality among users, we only require (3) to be satisfied for  $j > i$  — to the left of  $\mathbf{A}_j$ . Hence user  $j$  must select their transmit weight matrix to compensate for the preceding interferers ( $i = 1, \dots, j-1$ ) but must not interfere with those users. The algorithm as described in the references uses maximal-ratio combining at the receiver.

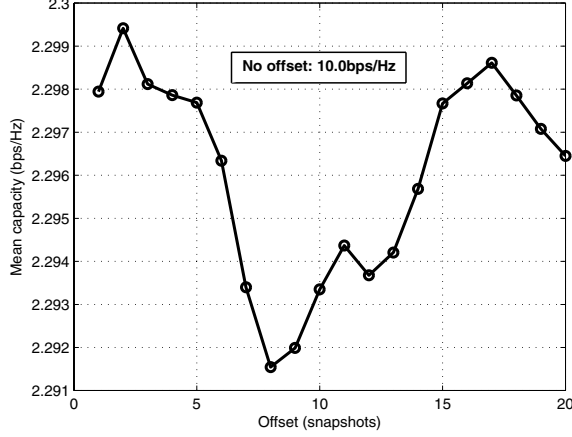
## IV. RESULTS

To allow a fair comparison among the schemes, we will always choose the same positions for the wanted user and the interferers, although this distinction only really exists for the NE. The wanted user will be at position no. 15 and the interferers at nos. 16, 17 and 18. Each user's channels will be normalized as described in Section II-A with each user received at a mean SNR of 20dB. Our performance metric will be the mean total system capacity across all users. We

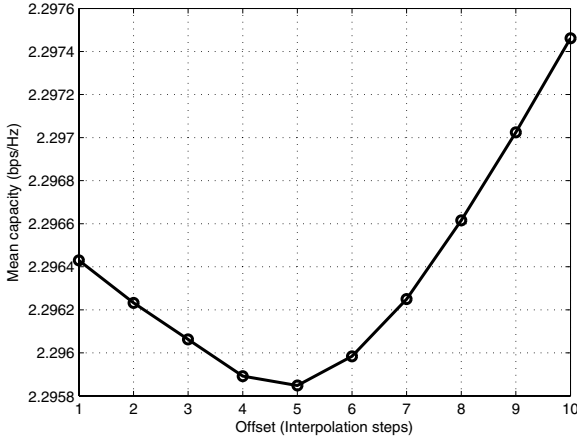
will restrict ourselves to considering a system in which  $n_T = 4$  for all users, and  $n_R = 4$  also. In light of (4), for the diagonalization-based schemes we will limit each user to a single stream, allowing water-filling to select that stream. Until Section IV-D, the thresholds for convergence will be a change of 1% or less in the total system capacity for the NE and all off-diagonal elements in (3) to be less than  $10^{-4}$  for BD.

### A. Snapshot-based delay

We begin by examining the losses on the coarser time-scale of the actual measured data, where each snapshot is approximately 10ms long. Figs. 2(a), 3(a) and 4(a) show how the capacity of the NE, BD and SD algorithms respectively varies as a function of offset. It is seen that the NE is easily the most robust to such delays. Its capacity does degrade noticeably, from about 22bps/Hz with no delay, but only as far as about 17bps/Hz, or approximately 23%. The block-diagonalization algorithm, on the other hand, is severely damaged even at the shortest time-offset available here (but



(a) Snapshot-by-snapshot capacity.



(b) Capacity at interpolated points between snapshots.

Fig. 4. Capacity of SD under channel offset. Capacity with no offset is 30.6bps/Hz. SNR = 20dB,  $n_T = n_R = 4$  for all users. Capacity with no offset is 10.0bps/Hz.

see below). The total capacity falls from about 30bps/Hz with instantaneous application of the transmit/receive weights to just 9bps/Hz at a one snapshot delay — a fall of 70%. It continues to decline sharply in the succeeding snapshots also. The successive diagonalization algorithm falls from about 10bps/Hz with no delays to about 2.3bps/Hz (falling by 77%) with a one snapshot delay. It then shows more or less random fluctuations at increased time-offsets. Note, however, the *extremely* expanded scale used in Fig. 4(a); on the scale of either of the other two schemes, these fluctuations would not be visible. For both BD and SD, these capacities would render the system unacceptable given the complexity of mulituser MIMO implementation — the per user capacities are approximately 2.3bps/Hz/user and 0.6bps/Hz/user respectively.

### B. Interpolation-based delay

By using the technique described in Section II-B, Figs. 2(b), 3(b) and 4(b) show each algorithm's capacity at ten equally-

spaced steps between zero and one snapshot of delay. Fig. 2(b) shows that the NE decreases at a steady, nearly linear, rate at this finer time-resolution, suggesting that delays on the order of a few milliseconds are unlikely to damage the algorithm's performance significantly. This conclusion is largely unchanged from that viewed at the coarser time-resolution of the original measured data where the loss at small offsets is also small. Block-diagonalization, on the other hand, shows rather different characteristics between Figs. 3(a) and 3(b). The results using interpolated channels show that the loss of channel capacity decreases significantly as the offset shortens. For delays in the region of a few milliseconds, the capacity falls by about 5–10bps/Hz from its value with instantaneous application of the transmit/receive weights and so BD can offer useful per-user capacities in this region. This is still very much more than the NE's loss, however, even at large time offsets. This is because BD depends on full orthogonality between users by design and the loss of orthogonality the time offset entails causes interference to reappear in evidently significant quantity. The NE, on the other hand, does not depend on orthogonality in the first place. It arranges instead to best tolerate the interference that is present and a disturbance from that arrangement is less severe.

In Fig. 4(b), we see that, even at the smallest time-offset we have available here, the capacity of SD is still degraded to the same level as at much larger offsets. The apparent 'minimum' capacity is an artefact of the sharp variations that are possible between one snapshot offset and the next, as seen in Fig. 4(a). In any case, the scale of the variation is very small indeed in comparison to the actual capacity values that are achieved and in comparison to the variations seen in the other two algorithms. Clearly, to recover a capacity approaching that of the ideal case, the computation and feedback delay for SD to operate usefully must be substantially under 1ms.

### C. Characterization of the losses

The curves in Figs. 2(b) and 3(b) may be conveniently characterized so that an analytical description of the behaviour is available which, in parameterized form, can be applied to other, similar, environments. We will not try to characterize Fig. 4(b) since the variations are extremely small.

By inspection, we see that Fig. 2(b) is nearly linear over the offset ranges 1–4 and 4–10. A least-squares linear regression over these two ranges gives:

$$C_{NE}(N_{\text{off}}) = \begin{cases} -0.2N_{\text{off}} + 22.7 : 1 \leq N_{\text{off}} \leq 4 \\ -0.4N_{\text{off}} + 23.6 : 4 \leq N_{\text{off}} \leq 10 \end{cases} \quad (5)$$

with  $N_{\text{off}}$  the number of offset units, with one unit being 0.9ms as before. These two straight lines are shown in Fig. 2(b) and clearly have a good fit. The two predictions at  $N_{\text{off}} = 4$  are very nearly identical.

For BD, we will try fitting an exponential decay of the form  $C = Ae^{-\lambda N_{\text{off}}}$ , again by least-squares regression over the full span of  $N_{\text{off}}$ . This gives

$$C_{BD}(t) = 31.1e^{-0.11N_{\text{off}}} \quad (6)$$

TABLE I  
EFFECT OF CHANGING CONVERGENCE THRESHOLD FOR THE NE

Convergence (% capacity)	1%	2%	3%	5%	10%
Mean iterations	4.5	4.2	4.0	3.8	3.4
Mean capacity (bps/Hz)	22.5	22.5	22.4	22.3	22.1

TABLE II  
EFFECT OF CHANGING CONVERGENCE THRESHOLD FOR BD

Convergence, eq. (3)	$10^{-6}$	$10^{-5}$	$10^{-4}$	$10^{-3}$	$10^{-2}$
Mean iterations	6.4	5.8	5.0	4.2	3.3
Mean capacity (bps/Hz)	30.4	30.3	30.2	30.1	29.4

This is a good fit; the mean absolute error is 1.9%. Thus we may conclude that the loss in capacity for BD at small time-offsets is approximately exponential; a convenient characterization.

#### D. Convergence -vs- delay

For the two iterative algorithms, some threshold must be set to determine when convergence has been ‘reached’ and the algorithm can be terminated. If a less stringent threshold is set, the algorithm can be terminated sooner in exchange for a poorer performance. However, this earlier termination will mean that the capacity loss due to delay in applying the weights is less and so in this part we explore this tradeoff.

Tables I and II summarise the effect on mean capacity with zero delay. It is immediately apparent from the tables that there is negligibly little capacity loss for considerably relaxing the convergence criteria. However, there are appreciable gains in the number of iterations required, especially with BD, where the mean iterative delay falls by nearly half over the range shown here. With the approximately exponential decay of Fig. 3(b), at delays of a few milliseconds this could give noticeable capacity improvements; certainly the gain in capacity from reduced delay would be substantially larger than the loss from weaker convergence. It is therefore important to operate the BD such that the algorithm can be terminated as soon as possible. The NE, with its much smaller losses to due to delay shows that a reduction by a quarter in the iteration time, as found here, would yield capacity gains of about the same size as the loss from weaker convergence. There is thus little overall change in the capacity, and the algorithm can be operated at the weaker convergence levels to reduce complexity and save power.

#### V. CONCLUSIONS

We have examined the effect of time-delay in the application of transmit/receive weights for three multiuser MIMO algorithms. We found that the NE is significantly more robust to temporal evolution in the channel than either BD or SD and that SD suffers considerably more than BD. At the time-resolution of the snapshots in the measured data (about 10ms), we found that BD and SD are both reduced to capacities that would be unlikely to justify the use of these algorithms in practise as they yielded capacities per user.

By interpolating between measured snapshots, we also examined the effect of delays between about 1ms and 10ms. This showed that both NE and BD can offer capacities much closer to the ideal case of zero delay at smaller offsets. However, even with a delay of 0.9ms, the capacity of SD showed no improvement over that at much longer delays. We also found simple analytical characterizations of the losses at the finer time-resolution by means of least-squares regression.

For the two iterative algorithms, we explored the tradeoff between weaker convergence (lower capacity) and lower delay (higher capacity). In both cases, we found that there is minimal capacity loss caused by allowing weaker convergence at zero delay. For the NE at small to medium delays, we found that the loss for weaker convergence approximately matched the gain for less delay whilst for BD, the gains for reducing delay were higher than the convergence losses. Both algorithms can therefore be operated at the weakest acceptable convergence thereby reducing the complexity and power consumption of a given implementation.

#### ACKNOWLEDGEMENTS

The authors would like to thank the UK EPSRC and Philips Research Laboratories for supporting this work.

#### REFERENCES

- [1] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, “Capacity limits of MIMO channels,” *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–701, June 2003, invited paper.
- [2] M. W. Webb, M. Hunukumbure, M. A. Beach, and A. R. Nix, “Evaluation of the capacity of multi-user MIMO schemes using feedback in a small urban cell,” in *Proc. 6th IEE International Conference on 3G and Beyond*, London, UK, Nov. 2005.
- [3] S. Ye and R. S. Blum, “Optimized signaling for MIMO interference systems with feedback,” *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.
- [4] Q. H. Spencer and A. L. Swindlehurst, “A hybrid approach to spatial multiplexing in multiuser MIMO downlinks,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2, pp. 236–247, 2004.
- [5] Q. H. Spencer and M. Haardt, “Capacity and downlink transmission algorithms for a multi-user MIMO channel,” in *Proc. 36th Asilomar Conference on Signals, Systems and Computers*, vol. 2, Nov. 2002, pp. 1384–1388.
- [6] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels,” *IEEE Trans. Signal Processing*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
- [7] Z. G. Pan, K. K. Wong, and T. S. Ng, “MIMO antenna system for multi-user multi-stream orthogonal space division multiplexing,” in *Proc. IEEE International Conference on Communications 2003*, vol. 5, Anchorage, Alaska, May 2003, pp. 3220–3224.
- [8] M. Hunukumbure, M. Beach, C. Williams, and S. E. Foo, “MIMO capacity limits in outdoor 3g systems,” in *Proc. 13th IST Mobile and Wireless Communications Summit*, Lyon, France, June 2004.
- [9] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, “Link-optimal space-time processing with multiple transmit and receive antennas,” *IEEE Commun. Lett.*, vol. 5, no. 3, pp. 85–87, Mar. 2001.
- [10] B. Farhang-Boroujeny, Q. Spencer, and L. Swindlehurst, “Layering techniques for space-time communication in multi-user networks,” in *Proc. 58th IEEE Vehicular Technology Conference, Fall 2003*, vol. 2, Orlando, Florida, Oct. 2003, pp. 1339–1343.
- [11] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, “Simplified processing for high spectral efficiency wireless communication employing multi-element arrays,” *IEEE J. Select. Areas Commun.*, vol. 17, no. 11, pp. 1841–1852, Nov. 1999.